

Resetting the pitch-analysis system: 1. Effects of rise times of tones in noise backgrounds or of harmonics in a complex tone

ALBERT S. BREGMAN, PIERRE AHAD, JEAN KIM, and LEANNE MELNERICH
McGill University, Montreal, Quebec, Canada

The question of whether sudden increases in the amplitude of pure-tone components would perceptually isolate them from a more complex spectrum was investigated in two experiments. In Experiment 1, a 3.5-sec noise was played as a masker. During the noise, two pure-tone components of different frequencies appeared in succession. Subjects were asked to judge whether the pitch sequence went up or down. The rise time of these components had only a small and inconsistent effect on discrimination. In Experiment 2, the 3.5-sec background signal was a complex tone. The amplitudes of two of its components were incremented in succession. Again, subjects judged whether the pitch pattern went up or down. This time there was a sizable, monotonic effect of the rise time of the increments, with more rapid increments leading to better discrimination. The difference between the two results is interpreted in terms of the auditory system's response to changing and unchanging signals and the role of its "sudden-change" responses in attracting perceptual processing to certain spectral regions.

Recently, Bregman, Ahad, and Kim (1992) studied the effects of suddenness of onset or offset on the clarity of pitches of individual pure tones that appeared in a 1-sec, 4-tone cluster of co-occurring tones. Two of their stimuli are shown in Figure 1. The corresponding upper and lower panels are different descriptions of the same condition. The upper panels show the temporal description, and the lower ones show the spectral description (except that onsets are emphasized by short vertical lines). The three frequencies that Bregman et al. (1992) used are shown: X = 800 Hz, H = 850 Hz, and L = 750 Hz. Note that they are all within a critical band. All components in the cluster on the left of Figure 1 have a rapid onset (10 msec), and their onset order is XH LX. The components of the cluster on the right have a gradual onset (640 msec), and their onset order is XLHX. The four tones have linear onset and offset amplitude envelopes with no steady states. In the Bregman et al. (1992) experiment, tones came on at a fixed stimulus onset asynchrony (SOA). The later-starting tones rose in amplitude only until they met the falling envelope of the first tone and then declined in amplitude along with it, as shown in the upper panels. It was found that faster rise times (in the range of 40 to 640 msec, as measured on the first tone

of the cluster) increased the accuracy of the detection of the order of onsets. A similar sort of effect was found when the clusters were digitally reversed. The stimulus for the resulting pattern is simply the mirror image of Figure 1. In this case, all the tones in the cluster rose at the same time but decayed asynchronously. More sudden offsets facilitated the discrimination of the order of offset. However, it was very much easier to detect onset order than offset order. With more sudden onsets or offsets, subjects also judged that they could hear a greater number of distinct sounds, and that the pitches were clearer. When asked whether the asynchrony was at the beginning or end of the stimulus, they were more accurate with the more rapid amplitude transitions. All judgments indicated that onsets were clearer than offsets.

Bregman et al. (1992) related their results to those of Pastore, Harris, and Kaplan (1982), who had studied the discrimination of onset order using a simpler stimulus. Two sine-wave tones, 1,650 and 2,350 Hz, started asynchronously (both employing the same rise time), held steady states for some duration and ended synchronously. Listeners were asked to determine whether the onset order was low-high or high-low. They found, as Bregman et al. (1992) did, that the discrimination varied with the rate of onset of the tones, more abrupt onsets being easier. The Bregman et al. (1992) experiments extended this finding to clusters of four tones in which the order of onset of the tones turning on second and third had to be discriminated. This meant that the discrimination could not be based on either the first or the last tone. They also extended the Pastore et al. results by showing that the tones in the cluster could all be within a critical band.

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the FCAR program of Quebec. J. Kim and L. Melnerich are affiliated with the School of Communication Disorders at McGill. Reprints are available from the first author at the Psychology Department, McGill University, 1205 Dr. Penfield Ave., Montreal, PQ, Canada H3A 1B1.

—Accepted by previous editor, Charles W. Eriksen

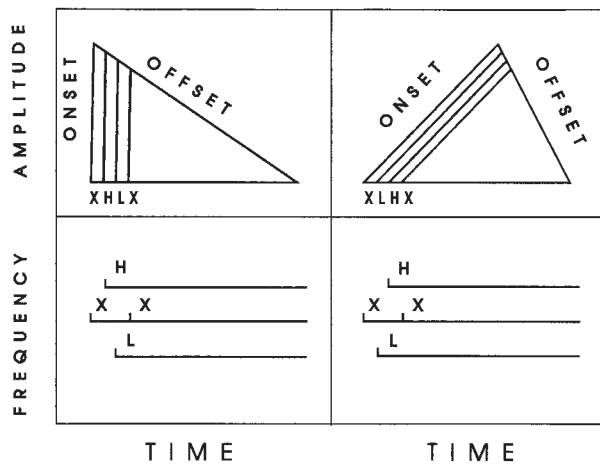


Figure 1. Diagram of two of the 1-sec tone clusters used by Bregman et al. (1992). The top boxes show amplitude envelopes of components superimposed; the lower boxes provide a spectrographic representation. Rise time of the first component is 10 msec in the left panel and 640 msec in the right panel; onset asynchrony is 60 msec. Onset orders are XHLX (left) and XLHX (right). X = 800 Hz, H = 850 Hz, L = 750 Hz.

Bregman et al. (1992) also related their observations on four-tone clusters to the research of Kubovy (1981), who reported a number of experiments in which a chord of pure tones was presented. When they were all on at the same time and steady in amplitude, there was no perceptual domination by the pitch of any one. However, when one of the tones was lowered in amplitude for 0.1 sec and then suddenly restored to its original value, its pitch dominated the experience of the subject. Kubovy (1976) called the phenomenon the “onset-segregation” effect and attributed it to the fact that there were specific neural responses to the onset of a sound. Bregman et al. (1992) proposed that the name “onset segregation” was too restricted because they had also found an offset effect that was rate-dependent. They suggested that it was more appropriate to call it a “sudden-change” segregation effect. They concluded that the auditory system might use neural onset and offset responses to reset itself and carry out new analyses at points of sudden amplitude change.

There was one difficulty with this interpretation. Figure 1 shows a fast-rising cluster of tones on the left and a slow-rising one on the right. Consider the cluster shown on the left. At the instant at which the second tone, H, starts, the first tone, X, is already decreasing from its maximum amplitude; that is, the difference in amplitude between X and H is very large. Now, looking over at the cluster shown on the right, we can see that at the instant at which the second tone begins, the first tone is still quite low in amplitude. The large difference between the two clusters, in the relative amplitudes of their first two tones, continues until the second tone reaches maximum amplitude. The same reasoning applies to all consecutive pairs of tones. Only in the fast-rising clus-

ter does an earlier tone reach its maximum amplitude before the next one has risen very high in amplitude. This is an inevitable consequence of the fixed SOA and the linear rising envelopes. The described difference between the suddenly and the slowly changing envelopes could conceivably have been responsible for the greater clarity of the pitches in the rapid-onset cluster.

The experiments that will be reported here investigated the rapid-onset effect without producing this problem. Rather than being overlapped in time, two pure tones were presented in succession. Listeners were asked to report whether the pitch pattern went up or down. The two tones were presented in a background of white noise. As in the experiments of Bregman et al. (1992), the tones had linearly rising and falling amplitude envelopes and no steady states, and the rise times of the tones were varied. It was anticipated that more rapid rise times would attract processing to the frequency band of the tone and lead to more accurate discrimination. This did not occur.

EXPERIMENT 1

Method

Procedure. On each trial, subjects heard a warning tone, followed, 1-sec later, by a noise lasting for 3.5 sec. During the noise, two 1-sec-long tones of different frequencies were presented in succession. The first tone was presented 0.5 sec after the noise onset, and the second tone was presented 0.5 sec after the end of the first. After the sounds, the subjects were asked whether the pitch pattern had gone up or down and to rate their level of confidence. They did this by pressing one of six keys on a keyboard, numbered from 1 to 6. The left side of the scale, numbers 1 to 3, was labeled “up,” and the right side, numbers 4 to 6, was labeled “down.” The extremes of the scale indicated certainty; the numbers moving toward the center of the scale indicated increasing uncertainty. The subjects could take as long as necessary to record their responses, after which there was a 1-sec silence, followed by the next trial.

The experiment was divided into six sessions. The first was a training session in which the rise time was progressively increased in duration across conditions (i.e., more sudden onsets first). The amplitudes of the tones were much higher in this training session than they were in the experiment proper. Also, before each trial, the listeners were informed as to whether the pattern went up or down. In the second session, the 12 training conditions were presented in a random order as a pretest. In order to proceed to the experiment proper, the subjects had to be correct on eight conditions. The other four sessions were test sessions of 72 trials each. There was a rest period of 2 to 5 min between test sessions.

Stimuli. The 3.5-sec masking noise was digitally bandpass filtered with upper and lower cutoff frequencies of 2500 and 500 Hz, respectively. To ensure that some fixed properties of the noise samples used for a particular condition were not responsible for the results of that condition and could not serve as a cue to the identity of the condition, each condition was synthesized with four independent noise samples. Since there were 72 different experimental conditions, 288 different samples of noise were used in all.

There were six different rise times for the 1-sec tones in the test trials: 30, 90, 270, 730, 910, and 970 msec. Notice that because the tones had linear rise and fall envelopes and no steady states, the last three envelopes were the mirror images of the first three. Each envelope had the same total energy as its mirror twin.

The level reached at the maximum of the triangular tone envelope is expressed in decibels relative to the noise level, that is, in signal-to-noise (S/N) level, where both tone and noise are measured in SPL (sound pressure level) at A weighting. The noise was always presented at 68 dB, and there were three maximum S/N levels for tones: 1, 2, and 3 dB. These values were chosen after pretesting to yield a level of discrimination which we thought would produce a sufficient number of errors for the detection of differences among the conditions. Within a given trial, both tones had the same S/N level. Three different tone frequencies were used: a high tone (H) at 2000 Hz, a medium tone (M) at 1500 Hz, and a low tone (L) at 1000 Hz. These appeared in ascending pairs, LM or MH, and in descending pairs, HM or ML.

The tones used in training were at 12 dB S/N, and their three rise times were 1, 10, and 150 msec. None of these values appeared among the signals of the actual test sessions. The warning tone for each trial in all sessions was at 500 Hz and lasted 150 msec.

Apparatus. The sounds, digitally synthesized on an IBM-compatible 386 computer using MITSYN software (Henke, 1981), were presented via 16-bit D/A converters (Data Translation 2823) at 16,000 samples per second. Signals were low-pass filtered at 8 kHz by a Rockland Model 852 filter (Butterworth setting) with 48 dB/octave roll-off. The sounds were presented diotically over Sony Model NRV7 headphones in an Industrial Acoustics single-wall test chamber. Sound pressure levels were measured at fast A weighting using a flat-plate coupler.

Subjects. Twenty people, with a mean age of 24.9 years and a range of 21 to 51 years, were paid to serve as subjects. None reported problems in their hearing. None were discarded and no screening tests were administered.

Results

There were 288 experimental conditions obtained by combining 6 values of rise time, 3 S/N levels, 4 frequency patterns (2 up and 2 down), and 4 different samples of background noise within each condition. The results were combined across variations in background noise samples, since these variations were introduced only so that we could average out the peculiarities of particular noise samples.

In our treatment of the data, we assume that the subject registers the stimuli on an underlying one-dimensional scale of "perceived direction," ranging from up to down, with an uncertain area in the middle. So far, we are making the same assumptions as would be necessary if we were to proceed to a signal detection (d') analysis by assuming that the subject puts a boundary on this dimension and answers "up" or "down" for the two ranges separated by the boundary. However, our method makes a different assumption, namely that the subjects' overt responses on the one-dimensional rating scale are monotonically related to the underlying dimension. We compare the scores for the ascending and descending patterns within each condition. If the mean ratings for ascending and descending patterns are the same, we interpret that as a lack of discrimination. If they differ in the right direction (descending patterns giving scores closer to the "certain-down" end), we take that as a sign of discrimination because the response is covarying with the stimulus. We assume that the greater the difference in the average ratings assigned to ascending and descending patterns, the greater the discrimination. This set of assumptions makes it possible to assign a discrimination

score for each subject within each combination of rise time, size of amplitude increment, and session. Because of the small number of trials (18) within such cells, analysis via signal detection theory to yield d' scores is not advisable. However, our discrimination score, based on means, does not require a large number of values per cell.

Accordingly, the rating scale scores were turned into a discrimination score, C (for covariation of the response with the stimulus), by the formula $C = [(D1 + D2) - (U1 + U2)]/2$, where $D1$ = rating for the HM descending frequency pattern, $D2$ = rating for the ML descending pattern, $U1$ = rating for the LM ascending pattern, and $U2$ = rating for the MH ascending pattern.

The formula gives scores ranging from +5 for perfect discrimination (or covariation), through 0 for random responding, to -5 for systematically reversed ratings for up and down.

Table 1 gives the raw rating-scale score for the six rise times and the three signal-to-noise ratios for both "ascending" and "descending" stimulus patterns. It also gives the derived C score for each condition. Within the limits of round-off error, the difference between each successive pair of raw scores (descending minus ascending) yields the corresponding C score. Both univariate and multivariate tests were carried out on all variables. Only the univariate ones will be reported unless one is significant and the other is not. Figure 2 shows the main effect of rise time. The effect of rise time was statistically significant [$F(5,95) = 12.2, p < .0001$]. There was, however, only a small and irregular decline in discrimination with longer rise times, with a slight rise again at the longest rise time. Note, however, that the longest rise time (970 msec) is also the shortest offset time (30 msec) because the tones lasted 1 sec and there was no steady state. The small increase in discrimination

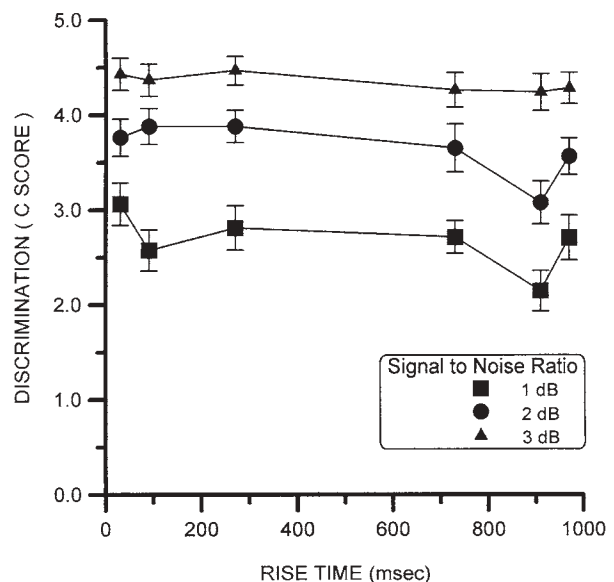


Figure 2. Rise-time results for Experiment 1. Maximum possible score is 5.

Table 1
Experiment 1. Mean Raw Scores and C Scores for Tones
in a Noise Background (Number of Cases: 20)

Rise Time (msec)	S/N (dB)	Actual Sequence	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>C</i> Score*
30	1	Ascending	1.98	.168	.752	3.06
		Descending	5.04	.093	.418	
30	2	Ascending	1.66	.140	.628	3.76
		Descending	5.43	.109	.487	
30	3	Ascending	1.40	.127	.567	4.43
		Descending	5.83	.074	.330	
90	1	Ascending	2.28	.137	.612	2.58
		Descending	4.86	.118	.528	
90	2	Ascending	1.69	.119	.533	3.88
		Descending	5.57	.107	.479	
90	3	Ascending	1.39	.112	.501	4.37
		Descending	5.76	.086	.384	
270	1	Ascending	2.03	.140	.625	2.81
		Descending	4.84	.130	.582	
270	2	Ascending	1.70	.122	.546	3.88
		Descending	5.58	.077	.342	
270	3	Ascending	1.33	.100	.445	4.47
		Descending	5.79	.069	.307	
730	1	Ascending	2.15	.117	.525	2.71
		Descending	4.86	.111	.497	
730	2	Ascending	1.64	.124	.556	3.65
		Descending	5.29	.141	.631	
730	3	Ascending	1.39	.098	.438	4.26
		Descending	5.65	.103	.460	
910	1	Ascending	2.39	.131	.585	2.14
		Descending	4.53	.132	.592	
910	2	Ascending	1.98	.141	.632	3.08
		Descending	5.05	.112	.501	
910	3	Ascending	1.43	.132	.591	4.24
		Descending	5.66	.090	.402	
970	1	Ascending	2.26	.161	.721	2.71
		Descending	4.96	.123	.549	
970	2	Ascending	1.89	.120	.538	3.56
		Descending	5.45	.084	.377	
970	3	Ascending	1.46	.105	.471	4.28
		Descending	5.74	.087	.390	

Note—For the mean raw scores, higher values indicate stronger judgments of “downness.” *Within the limits of round-off error, *C* scores are the difference between the descending and ascending scores for the same rise-time-by-increment condition.

may therefore be due to the “abruptness-of-offset” effect found by Bregman et al. (1992).

There was a clear effect of S/N level, with higher levels producing better discrimination [$F(2,38) = 200.5$, $p < .0001$]. The mean *C* scores were 2.7, 3.6, and 4.3 for S/N levels of 1, 2, and 3 dB, respectively. There was no significant interaction with rise time.

Discussion

Although a small and irregular decline was observed for discrimination as a function of rise time, the influence of rise time was much smaller than that reported by

Bregman et al. (1992) in clusters of tones. The results were quite significant statistically, but this was due only to the statistical power of the analysis, not to the size of the effect.

One important difference between the present experiment and those of Bregman et al. (1992) is the fact that in the present experiment it was the presence of noise that created the difficulty in detecting the pitch pattern, whereas in the earlier ones it was the presence of the other tones in the cluster. Experiment 2 tested the possibility that the large effects of the earlier experiments involved segregation from a tonal context and not purely an increase of resistance to masking in general.

EXPERIMENT 2

In this experiment, we tried to set up conditions in which the problem was not detection of a masked signal but segregation from a tonal context. It is possible that the classical experiments on “masking” and the experiments on “fusion” (of components of mixtures) may involve different underlying processes (Bregman, 1990, pp. 314–325). The total signal that was presented to subjects (including both background and target components) consisted of a fundamental at 500 Hz and its next four harmonics. These components were present at all times. The amplitudes of an adjacent pair of the frequencies, 1000, 1500, and 2000 Hz, were incremented in succession, and the subjects had to judge whether the pitch pattern went up or down. We varied the suddenness of onset of the increments and the magnitude of the increments. A diagram of the harmonic tone and an increment in two of its harmonics is shown in Figure 3 (the size of the increment is not drawn to scale).

Method

Procedure. The procedure was the same as that in Experiment 1 except that, instead of a noise background, we presented a complex tone consisting of the first five harmonics of a 500-Hz fundamen-

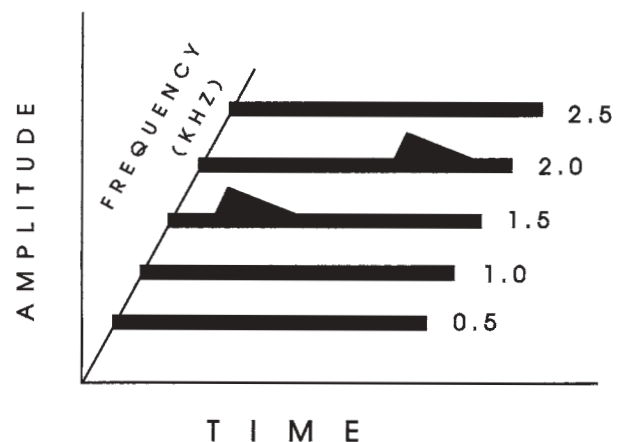


Figure 3. Diagram of the stimuli of Experiment 2, showing an MH frequency pattern for a moderately slow rise time. The size of the increment is exaggerated for clarity.

tal, which lasted for 3.5 sec. At a point 0.5 sec after the onset of the complex tone, one of its harmonics was incremented in amplitude and then fell to its initial value; the total rise-and-fall change was completed in 1 sec. Then, 0.5 sec after this event, a different harmonic underwent an identical change, in terms of both rise time and amplitude. After the complex sound finished, the subjects were asked whether the pitch pattern of increments had gone up or down, and to rate their level of confidence, as in Experiment 1.

The experiment was divided into six sessions. The first two were the training and test sessions, which followed the same procedures as in Experiment 1. The other four sessions were test sessions of 72 trials each.

Stimuli. The envelope of the rise and fall of amplitude of the increment had the same form as the rise-and-fall envelope of the tones of Experiment 1. That is, they rose linearly from their resting values to a maximum and then immediately fell linearly back to their resting values. The six rise times were the same as those employed in Experiment 1. The level reached at the maximum of the triangular tone envelope is expressed in decibels relative to the steady-state level before and after the increment. The overall intensity of the complex tone in its steady state was 65 dB, A weighting. There were three increment levels: 1, 3, and 6 dB. These increments were chosen empirically, after pretesting 10 subjects, as being difficult enough for a high level of correctness not to obscure the effects of rise time on discrimination. Within a given trial, both tones had the same increment level. Values from Experiment 1 were used for the following: (1) the frequencies of the ascending and descending patterns, (2) the onset times and frequencies of the tones used in training (but they now had a 12-dB increment), and (3) the warning tone.

Apparatus. The apparatus was the same as in Experiment 1.

Subjects. Nineteen people were paid to serve as subjects. Their mean age was 22.7 years, with a range of 21 to 27 years. None were discarded and no screening tests were administered.

Results

A discrimination score, *C*, ranging from +5 (perfect discrimination) through 0 (random responding) to -5 (systematically reversed ratings), was derived as in Experiment 1. Table 2 shows the raw rating-scale scores. High values signify the "downness" of the rating. The derived *C* scores are also shown. Within the limits of round-off error, the difference between each successive pair of raw scores (descending minus ascending) yields the corresponding *C* score.

Both univariate and multivariate tests were carried out on the variables of rise time, amplitude of increment, and session number. Only the univariate analyses will be reported unless one is significant and the other not. Figure 4 shows the effect of rise time separately for the three different increment amplitudes. The main effect of rise time was statistically significant [$F(5,54) = 36.5, p < .0001$]. The means for the six rise times, averaged over increment amplitudes, were 3.7, 3.5, 2.9, 2.2, 2.1, and 2.0. These means fall monotonically with rise time, and the decline is sizable in magnitude, covering about a third of the range between the maximum possible score of 5 and the random value of 0. The significant effect does not arise simply from the fact that the means of the first two rise times are higher than the others. There is also a significant difference between the 270- and 730-msec rise times [$F(1,18) = 20.3, p < .0005$]. The

Table 2
Experiment 2. Mean Raw Scores and *C* Scores, for Increments in Harmonically Related Tones (Number of Cases: 19)

Rise Time (msec)	Incr. (dB)	Actual Sequence	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>C</i> Score*
30	1	Ascending	1.72	.181	.790	3.64
		Descending	5.37	.140	.610	
30	3	Ascending	1.78	.201	.875	3.59
		Descending	5.38	.151	.658	
30	6	Ascending	1.63	.183	.797	3.89
		Descending	5.52	.150	.653	
90	1	Ascending	1.99	.148	.645	3.32
		Descending	5.31	.110	.479	
90	3	Ascending	1.85	.183	.799	3.51
		Descending	5.36	.153	.667	
90	6	Ascending	1.72	.153	.669	3.59
		Descending	5.31	.161	.700	
270	1	Ascending	2.38	.191	.832	2.54
		Descending	4.93	.167	.729	
270	3	Ascending	2.28	.197	.858	2.86
		Descending	5.14	.145	.633	
270	6	Ascending	1.97	.155	.677	3.24
		Descending	5.20	.164	.713	
730	1	Ascending	2.64	.157	.686	1.93
		Descending	4.59	.127	.554	
730	3	Ascending	2.64	.145	.632	2.07
		Descending	4.70	.164	.715	
730	6	Ascending	2.25	.161	.702	2.47
		Descending	4.72	.170	.743	
910	1	Ascending	2.67	.149	.651	1.61
		Descending	4.28	.199	.867	
910	3	Ascending	2.44	.150	.653	1.96
		Descending	4.39	.169	.738	
910	6	Ascending	2.27	.157	.682	2.61
		Descending	4.89	.173	.756	
970	1	Ascending	2.64	.169	.736	1.72
		Descending	4.37	.199	.865	
970	3	Ascending	2.38	.148	.643	1.97
		Descending	4.36	.206	.897	
970	6	Ascending	2.14	.174	.757	2.37
		Descending	4.52	.202	.882	

Note—Within the limits of round-off error, *C* scores are the difference between the descending and ascending scores for the same rise-time-by-increment condition. *For the raw scores, higher values indicate stronger judgments of "downness."

differences among the three slowest rise times are not significant.

The rise-time effect is much larger and more regular than in Experiment 1. Note that for these means, there is no violation of the monotonic decrease with longer rise times, even at the longest one (which had the most abrupt offset). A final point is that although the signals for the first three and last three rise times were mirror images, and therefore matched for energy, there is no sign of a U-shaped trend in the data that would suggest that the results were due to total energy rather than rise time.

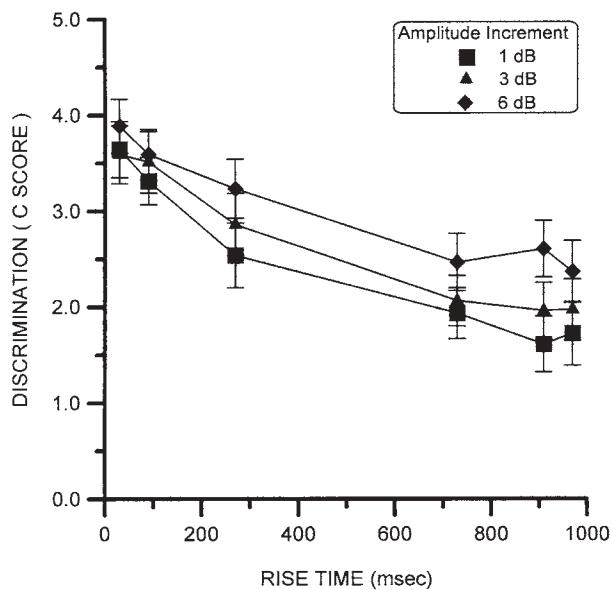


Figure 4. Rise time results for Experiment 2, plotted separately for the three different increment sizes (1, 3, and 6 dB). Maximum possible score is 5.

There was also a strong and significant effect of the size of the amplitude increment [$F(2,36) = 25.1, p < .0001$]. The means were 2.5, 2.7, and 3.0 for increments of 1, 3, and 6 dB, respectively.

The effect of test session (1 to 4) was significant [$F(3,54) = 10.6, p < .0001$]. The effect was of moderate size; the means were 2.4, 2.7, 2.8, and 3.0 for the four test sessions. Some degree of learning evidently took place.

There were no significant interactions among the three variables. The effect of rise time was found at all three levels of amplitude increment. It was notable that the rise-time effect occurred with even a 1-dB increment [$F(5,90) = 26.0, p < .0001$]. It is interesting to observe that, as shown in Figure 4, the discrimination scores for a given rise time fall within a relatively small range, especially at the 30- and 90-msec rise times.

GENERAL DISCUSSION

The results of Experiments 1 and 2 were quite different. When the target tones appeared in a background of noise, just above the noise level, the effects of rise time were neither large nor monotonic. When the mean discrimination scores for different rise times were compared (averaged across other variables), the range from lowest to highest was only 0.6, or 12% of the positive range of the scale. Although the results were very significant, this was due more to the power of the statistics than to the size of the results. The results were quite different from those of Bregman et al. (1992), which showed a strong effect of rise time, with a range of values covering about 33% of the positive range of the scale (recall that a mean of 0 corresponds to random perfor-

mance and that the negative range corresponds to systematically reversed judgments, which are rarely observed).

On the other hand, in Experiment 2, where the background of the incremented harmonic was other unchanging harmonics, discrimination fell monotonically as rise time increased and the range of means (averaged over other variables) was 1.7, or 34% of the positive range of the scale. The magnitude of this result is quite similar to that of the result of Bregman et al. (1992).

The difference between the present two experiments can be explained in the light of the theory proposed by Bregman et al. (1992) to account for their results. As we outlined in the introduction, they presented a cluster of pure tones, all within a critical band, having asynchronous onsets but synchronous offsets. They reported that when they listened to versions of their stimuli that were not asynchronous, they found it difficult to hear three different pitches in the mixture. Only when a single component changed in amplitude did its pitch become salient. This suggested to them that sudden amplitude changes, by activating "onset" responses of the nervous system, are capable of causing a "resetting" of the pitch-analysis mechanism. The new analysis is dominated by the components that have suddenly changed. Presumably other tonal components that are unchanging do not interfere with the computation of the new pitch, despite being in the same critical band, because they trigger no auditory "onset" responses at that instant. These ideas are similar to those of Kubovy (1981).

An extension of this explanation will handle the case of Experiment 2 quite well. In that experiment, each harmonic's increment occurred when the other harmonics were steady in amplitude. We assume that the "sudden-change" effect described in the introduction applies not only to "true" onsets or offsets (from or to silence) but to amplitude increments as well. It seems plausible that the amplitude increments of this experiment activated the same sort of onset responses in the auditory system that are seen with true onsets. If so, Bregman et al.'s (1992) explanation can be applied here; the auditory system's onset responses to particular frequencies do two things: they directly supply frequency information involved in the identification of the tone and also point to a narrow spectral region for enhanced analysis by other processes. Such a mechanism may be involved in a range of phenomena on "auditory enhancement," studied by Viemeister (1980), Summerfield, Haggard, Foster, and Gray (1984), and others, and on the "tonal aftereffect" studied earlier by Zwicker (1964).

We must, however, explain the weakness or non-existence of a rate-of-onset effect in Experiment 1. Our previous explanation (Bregman et al., 1992), which was applied to onsets in a cluster of pure-tone components, attributed the lack of interference from the other simultaneous components to the fact that at the moment that the target component changed, the other components remained steady. Therefore, the spectral region of the target tone was the only one that was changing. For this rea-

son, the competing components did not attract spectral processing. The situation is different in the case of noise. The spectrum of noise, by definition, is always changing. There is no stability, even in the very short run, against which the change in a target spectral region stands out. When the first author listened to the sudden and slow onsets in noise, he actually judged that he could hear the tones with the slow onsets best because they sounded smooth in an otherwise constantly rough background. Other subjects did not experience this effect.

The explanation requires that the only onset (or amplitude change) activity be at the target frequency (or frequencies, in the case of a complex target). Onsets too near in time to that of the target tone are harmful. It is well known that in simultaneous masking, a longer delay of the onset of the target relative to that of the masker will improve detection (the “overshoot” effect studied by Zwicker, 1965, Bacon & Moore, 1986, McFadden, 1989, Carlyon & White, 1992, and many others). This is probably related to the sudden-change effect, and can be related to Experiment 1 as follows: There may be two classes of masking resulting from the amplitude changes in a noise mask. One (“onset masking”) may result from the sudden amplitude increase when the mask is turned on. The second (“ongoing masking”) may result from the amplitude changes occurring throughout the noise. (The latter would have been the basis for the masking observed in Experiment 1.) If one assumes that onset masking (being a result of greater amplitude changes than are present in ongoing masking) creates more interference, then it is possible that a target may be at a high enough amplitude to escape the ongoing masking but not the onset masking. This would explain why it could be detected more easily if its onset came well after the onset of the mask.

In Bregman et al.’s (1992) experiment, subjects were able to hear individual onsets, spaced as little as 60 msec apart, and within the same critical band as other components. Even among the onsets that began when the previous tone was still rising (i.e., when the rise times were longer than the onset asynchrony), the tones with more abrupt onsets were heard better. This suggests two things: (1) The onset response is not present throughout the whole rise but is triggered discretely at a certain point in amplitude (or amplitude change); and (2) as long as these discrete responses are separated sufficiently in time, they will enable processing at a specific spectrotemporal position to be dominant for a brief period of time. Without these assumptions, when the onsets of tones overlapped, the neural “onset” responses would also overlap, and therefore could not improve the segregation of the tones.

In the present Experiment 2, there was no competition among amplitude changes near each other in time. This is the best situation for the use of onset information. In the experiments of Bregman et al. (1992), there was sufficient time between the changes to allow them not to completely interfere with one another. The worst situation was found in the present Experiment 1, where the

constant change inherent in noise, together with the small S/N levels used, did not allow the onset responses to be restricted to the regions of the targets. Therefore, processing could not be helped by the onset-response mechanism.

Finally, it is remarkable that, in Experiment 2, an increment of even a single decibel could be perceived almost as easily as one of 6 dB, especially with rise times less than 100 msec. The spread of spectral energy with rapid rise times could allow detection of the onset of an event, but detection was not all that was required in the up-down judgment task. The actual pitch had to be detected (to some degree of accuracy). In any case, the spectral spread at any rise time beyond 30 msec should have been negligible. Yet the falling function for the 1-dB curve in Figure 4 is not flat from 90 msec onward; it shows significant differences out to 730 msec. Therefore, the spectral spread in the signal cannot have been playing much of a role.

The physiological basis for sudden-onset effects is not presently known. However, some evidence points to a rather peripheral origin. Delgutte (1980) measured post-stimulus-time histograms of the responses of an auditory nerve fiber in the anesthetized cat to a noise burst with either a sudden (1-msec) or a gradual (40-msec) onset. He found that the initial (early) neural activity was much larger for the sudden onsets. Later, working with synthetic speech sounds, he and Kiang showed that the 10-msec rise in the consonant “ch” led to stronger initial activity in the auditory nerve than did the 70-msec rise in the consonant “sh” (Delgutte & Kiang, 1984). However, one should be cautious about drawing inferences from these findings about the origins of the onset effects in the present experiment because of the different ranges of onset times involved. Our *fastest* onset was 30 msec, and the range over which Figure 4 shows differences is at least 700 msec. Furthermore, while the abruptness of onset may well be *coded* at a low level, it is likely that it is not used for purposes of segregation of sounds until it reaches a higher level, at which a variety of cues (such as harmonicity, asynchrony of onset, resemblance to earlier sounds, different spatial origin, etc.) can be brought together.

The logical way to continue this research would be to ask what would happen if several components of a complex spectrum rose in synchrony. One might expect that the auditory system would tend to treat this subset as a distinct perceptual entity and to compute new global properties for it. Would the emergence of these global properties, distinct from the properties of the larger spectrum, depend on the rise time of the subset?

REFERENCES

- BACON, S. P., & MOORE, B. C. J. (1986). Temporal effects in simultaneous pure-tone masking: Effects of signal frequency, masker/signal frequency ratio, and masker level. *Hearing Research*, **23**, 257-266.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- BREGMAN, A. S., AHAD, P., & KIM, J. (1992). *Resetting the pitch-*

- analysis system: 2. Effects of sudden onsets and offsets in a cluster of overlapping tones.* Manuscript submitted for publication.
- CARLYON, R. P., & WHITE, L. J. (1992). Some experiments relating to the overshoot effect. *Auditory perception and physiology: Proceedings of the 9th International Symposium on Hearing, Carcans, France, June 1991.*
- DELGUTTE, B. (1980). Representation of speech-like sounds in the discharge patterns of auditory-nerve fibers. *Journal of the Acoustical Society of America*, **68**, 843-857.
- DELGUTTE, B., & KIANG, N. Y. S. (1984). Speech coding in the auditory nerve: IV. Sounds with consonant-like dynamic characteristics. *Journal of the Acoustical Society of America*, **75**, 897-907.
- HENKE, W. L. (1981). *MITSYN: A coherent family of command-level utilities for time signal processing* [Computer program]. Belmont, MA: Author.
- KUBOVY, M. (1976, November). *The sound of silence: A new pitch-segregation phenomenon.* Paper presented at the 17th Annual Meeting of the Psychonomic Society, St. Louis.
- KUBOVY, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 55-98). Hillsdale, NJ: Erlbaum.
- McFADDEN, D. (1989). Spectral differences in the ability of temporal gaps to reset the mechanisms underlying overshoot. *Journal of the Acoustical Society of America*, **85**, 254-261.
- PASTORE, R. E., HARRIS, L. B., & KAPLAN, J. K. (1982). Temporal order identification: Some parameter dependencies. *Journal of the Acoustical Society of America*, **71**, 430-436.
- SUMMERFIELD, Q., HAGGARD, M., FOSTER, J., & GRAY, S. (1984). Perceiving vowels from uniform spectra: Phonetic exploration of an auditory aftereffect. *Perception & Psychophysics*, **35**, 203-213.
- VIEMEISTER, N. F. (1980). Adaptation of masking. In G. van den Brink & F. A. Bilsen (Eds.), *Psychophysical, physiological, and behavioural studies in hearing.* Delft: Delft U.P.
- ZWICKER, E. (1964). Negative after-image in audition. *Journal of the Acoustical Society of America*, **36**, 2413-2415.
- ZWICKER, E. (1965). Temporal effects in simultaneous masking by white-noise bursts. *Journal of the Acoustical Society of America*, **37**, 653-663.

(Manuscript received August 10, 1992;
revision accepted for publication February 26, 1994.)